Amendments to the Specification

Please replace paragraphs [0019]-[0020] with the following replacement paragraph.

[0019] A first step **1** in the method illustrated in Figure 1, identifies one or more alignment domains in each sequence corresponding to a structure returned from a database query. An alignment domain refers to: 1) to a sequence domain that is evolutionarily conserved across related sequences, or 2) the sequence domain corresponding to a structural fold that is conserved across related structures. One method for annotating one more alignment domains in a sequence uses RPS-BLAST in combination with a database of annotated sequence domains or sequence profiles such as the Pfam Database, the SMART Database, or the COG database to search a sequence for domains annotated in these respective databases. Links to each of these databases and a link for downloading RPS-BLAST is available at on the internet ~~http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml~~. Cut-off e-values for identifying a sequence domain may vary depending upon the required confidence of a domain identification, but an exemplary, suitable, cut-off e-value ranges from 10^{-4} to 10^{-2}, with 10^{-3} preferred.

[0020] Alignment domains may also be based upon structure-structure alignment domains. If a sequence corresponding to a structure returned from a database query is in the Protein Data Bank, and annotated in the SCOP database, ~~http://scop.berkeley.edu/,~~ SCOP structural annotations may be used to annotate it. In addition to using domain databases, one or more sequence domains may be identified in a sequence (referred to as the query sequence) by a method comprising the steps of: 1) determining a plurality of related template sequences to the query sequence using a sequence alignment tool such as the various BLAST, Smith-Waterman or FASTA

algorithms and a large protein sequence database such as the NCBI Protein Sequence Database, ~~http://www.ncbi.nlm.nih.gov/~~; 2) identifying putative conserved domains in the template sequences based upon their alignments with the query sequence; 3) performing a multiple sequence alignment on the related template sequences using a multiple sequence alignment tool such as ClustalW, ~~http://www.ebi.ac.uk/clustalw/~~; 4) identifying one or more conserved domains from the multiple sequence alignment; 5) determining one or more domain profiles or Hidden Markov Models ("HMMs") of the domains identified in step 4); and 6) identifying an alignment domain with a profile or HMM determined in step 5). A domain profile may be determined from a domain identified in a multiple sequence alignment using PSI-BLAST, ~~http://www4.ncbi.nlm.nih.gov/Education/BLASTinfo/psi1.html~~. A Hidden Markov Model may be built from a multiple sequence alignment using HMMER~~, available for download at http://hmmer.wustl.edu/~~. Methods for determining putative conserved alignment domains based upon alignment data are within the capacity of one ordinarily skilled in the art. Sonnhammer, E.L.L., and Kahn, D., *Modular Arrangement of Proteins as Inferred from Analysis of Homology,* PROTEINS: Structure, Function and Genetics, 3:482-492 (1994) discusses the methods for identifying one or more putative conserved alignment domains from sequence alignment data. Sonnhammer, E.L.L., Eddy, S.R., and Durbin, R., Pfam: *A Comprehensive Database of Protein Domain Familes Based on Seed Alignments*, PROTEINS: Structure, Function and Genetics, 28:405-420 (1997) details how the alignment domains in the PFAM database are determined.

Please replace paragraph [0029] with the following replacement paragraph.

[0029] A sixth step **9**, receives one or more sequence selections made by the user using the means presented to the user for selecting a sequence thereby selecting **13**

their corresponding structures for subsequent processing. Protein structure processing refers to the processing of the structure coordinate files corresponding to the selected proteins. Exemplary processing includes but is not limited to, visualization of the protein structure or further proteomic analysis such as, fold identification, the identification of functional sites on the protein, virtual ligand screening or small molecule docking. As one ordinarily skilled in the art will appreciate, the methods according to the invention are agnostic as to the nature of the post selection protein structure processing. If a structure has been selected for visualization, its coordinate file may be processed with protein structure viewing software in order to display a three dimensional model of the protein to a user. Exemplary software for displaying protein structures includes but is not limited to, Rasmol~~, available for download at http://www.rasmol.org/~~, Cn3D ~~available for download at http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml~~, Molscript~~, available for download at,~~ http://www.avatar.se/molscript/, MolMol ~~available for download at http://www.mol.biol.ethz.ch/wuthrich/software/molmol/~~, and the Insight II software suite available from Accelrys, Inc., (San Diego, California). For those structure viewers that recognize PDB structure files, a protein structure file may be formatted accordingly. ~~See http://www.rcsb.org/pdb/docs/format/pdbguide2.2/guide2.2_frame.html.~~ For those current or future viewers that do not recognize PDB structure files, a script may be written, using either the native scripting features in a viewer, or in an external scripting language, to format a structure file in a format that is recognized by a particular viewer.

Please replace paragraph [0050] with the following replacement paragraph.

[0050] Programming for displaying protein structures based upon their structural coordinates **130**, as used herein, refers to machine code, that when executed by the

processor, displays protein structures to the user via the output device, **123**, based upon their structural coordinates. Exemplary software for displaying protein structures includes but is not limited to, Rasmol, ~~available for download at http://www.rasmol.org/,~~ Cn3D ~~available for download at~~ ~~http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml~~, Molscript~~, available for download at,~~ ~~http://www.avatar.se/molscript/~~, MolMol ~~available for download at~~ ~~http://www.mol.biol.ethz.ch/wuthrich/software/molmol/~~, and the Insight II software suite available from Accelrys, Inc., (San Diego, Ca).